UNITED STATES OF AMERICA
Federal Trade Commission
WASHINGTON, D.C. 20580

Commissioner Maureen K. Ohlhausen
Office of the Commissioner

Aug. 5, 2014

Via Electronic Filing
National Telecommunications and Information Administration
U.S. Department of Commerce
1401 Constitution Avenue N.W.
Washington, DC 20230

Re: Comments of Maureen K. Ohlhausen, Commissioner, Federal Trade Commission on Big Data, Consumer Privacy, and the Consumer Bill of Rights[1]

Dear National Telecommunications and Information Administration,

I write to provide my perspective on the policy issues for which you recently issued a request for comment.[2]

As society has integrated and adopted increasingly powerful computers and pervasive communications networks, we have created immense amounts of data. Data today has greater volume, variety, and velocity. Put more simply, there is a lot of data, it has many different forms, and it is created rapidly. This trend will continue as we move into the era of the Internet of Things, a universe of Internet-connected devices that will tremendously increase the amount of passive data collection. Tools that can pull useful insights from this flood of data have great potential to make our lives better and promise significant benefits for consumers, businesses, and

[1] The views expressed here are solely those of Commissioner Ohlhausen and are not intended to reflect the views of the Commission or any other Commissioner.

[2] National Telecommunications and Information Administration, Docket No. 140514424-4424-01, Big Data and Consumer Privacy in the Internet Economy, 79 Fed Reg. 32714 (June 6, 2014) (RFC).

government. Big data techniques will help us extract knowledge from data, and this knowledge will help us better understand ourselves and the world around us.

Although big data will offer many benefits, the potential uses of big data also raise concerns about consumer privacy. Big data is a tool; like all tools it has strengths and weaknesses. Keeping those strengths and weaknesses in perspective is important as government, industry, research institutions, and civil society work together to adapt our laws, guidelines, best practices, and customs to integrate this new technology. As we adapt to big data, I believe the FTC can serve an important role in protecting consumers while allowing innovation to thrive.

I. Keeping Big Data in Perspective

Forming an adequate, proportional policy response to big data technology requires a realistic and accurate view of that technology. Data is invaluable to help us understand and affect change to the world around us. However, data, even big data, isn't knowledge or wisdom. It can be misleading. Accurately understanding both the benefits of and shortcomings of big data technology is critically important to getting the most out of this technology.

Big data is not completely new. Large companies like Wal-Mart have been collecting and analyzing terabytes of consumer data since the early 1990s.[3] However, today the tools for big data analysis are cost-effective even for small companies. Additionally, there is a lot more data, of many different kinds, being produced today.[4] Back then, Wal-Mart was collecting data on what consumers purchased from its stores and batch processing it each night. Today, a big

---

[3] *See* Thomas Wailgum, *45 Years of Wal-Mart History: A Technology Time Line*, Oct. 17, 2007, http://www.cio.com/article/147005/45_Years_of_Wal_Mart_History_A_Technology_Time_Line.

[4] According to Intel, in 2015 the world will produce 8.1 zettabytes of data – 1,500 times more than all the data produced from the beginning of time until 2003. *See* Intel, Big Data 101 Video, *available at* http://www.intel.com/content/www/us/en/big-data/big-data-analytics-turning-big-data-into-intelligence-cmpg.html (last visited Aug. 5, 2014).

data project might combine such purchase data with shipping, inventory, and even traffic data to achieve just-in-time delivery of consumer products.

Benefits of Big Data. This change in tools and data sources has great potential to make citizens' lives better. As Professor Sinan Aral of New York University has explained, "'Revolutions in science have often been preceded by revolutions in measurement.'"[5] The promise is that big data techniques will extract from data knowledge that will help us better understand the world, similar to how the microscope's magnification of tiny things led to the germ theory of disease.

Obviously, we are already seeing benefits from the use of big data techniques. Amazon, e-Bay, Netflix, and many other online merchants use big data to generate customized user recommendations. Other companies use big data to predict commute times by aggregating millions of GPS signals, to identify potential causes of disease, and to detect and prevent credit card fraud. Scientists are using massive data sets and powerful analytic tools to make progress on the most difficult problems in the health sciences and hard sciences. For example, Kaiser Permanente used big data analysis to discover an increased chance of heart attack or cardiac death among Vioxx users as compared to users of a competing medication. Congress is also considering a bill that would modify the national child death case reporting system to collect additional data that may allow big data techniques to help reveal the root causes, rates, and trend of sudden unexpected infant and child deaths.[6]

In addition, many new uses are emerging, particularly because consumers are no longer simply data points to be researched. Today's consumers are themselves producers and users of

[5] THE ECONOMIST, *Data, Data Everywhere*, Feb. 25, 2010, *available at* http://www.economist.com/node/15557443.

[6] Sudden Unexpected Death Data Enhancement and Awareness Act, H.R. 669, 113th Cong. (2013).

big data, whether posting billions of photos on Facebook, using Bing's flight price predictors to make travel plans, or joining the self-quantification movement by wearing a FitBit Flex. As more of our everyday existence becomes measurable and recordable, the greater potential there is for big data to provide helpful insights.

Challenges in Applying Big Data. Big data can be misleading, however. There can be "signal problems," where the data set, huge as it may be, does not represent the real world. The oft-cited City of Boston's StreetBump mobile app is an example of both how such bias can exist in the data *and* the appropriate response to such problems. The city recognized that the crowd-sourced pothole-finding mobile phone app would identify more potholes in wealthy areas of the city simply because more residents in those neighborhoods used smartphones. But the city didn't throw out the project because of this potential bias in the data. Instead, they started with a pilot program used by city workers, allowing the city to identify and adjust for the skewed data.

Big data sets can also let us mislead ourselves. As Nate Silver explains in his book "The Signal and the Noise," our instinctual shortcut when we have too much data is to pick out the parts we like and ignore the remainder.[7] Big data is particularly vulnerable to the "multiple comparisons problem," in part because big data tools are very good at discovering statistical correlations between variables in complex data sets. Because there are many variables in the typical big data set, there are many potential relationships for a researcher to test. If a researcher explores a big data set without a particular question or theory in mind but instead simply tries enough comparisons between variables, they will often be able to find "statistically significant" correlations that do not reveal anything useful about *causation* in the real world.

---

[7] Nate Silver, *The Signal and the Noise: Why So Many Predictions Fail - But Some Don't* at 3 (2012).

By better understanding the limits of big data and emphasizing the need for human judgment in the use of such tools, policy makers can help tamp down the hype – both for and against – over big data. Policy makers can help create a healthier regulatory atmosphere by critically evaluating the claims of both the pop-science promoters of big data as a "magic bullet" solution and the naysayers who fear enormous consumer harm from all-knowing algorithms. A realistic understanding of big data's potential will help policy makers to identify and focus on actual harms to consumers, if they occur.

II. Privacy and Other Concerns

As recognized in the White House's Big Data Report, big data raises concerns about how current laws will protect consumers.[8] Of course, many types of big data research have nothing to do with individuals and do not raise these concerns.[9] However, some consumer and privacy advocates are concerned that consumers will suffer harm from other uses of big data. Such advocates are particularly uncomfortable about the implications of large, persistent data sets containing information on individual customers. These concerns generally fall into three categories. First, some of the concerns about big data apply to data more generally, and the FTC has been actively addressing these issues for years. Second, as both the Big Data Report and the RFC indicate,[10] big data does raise some genuinely new challenges, particularly about how we

---

[8] Executive Office of the President, *Big Data: Seizing Opportunities, Preserving Values* at 48 (May 2014) (Big Data Report), *available at* http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf.

[9] For example, particle physics and astrophysics research produces enormous amounts of complex data. *See* Leah Hesla, "Particle Physics Tames Big Data," SYMMETRY, Aug. 1, 2012, *available at* http://www.symmetrymagazine.org/article/august-2012/particle-physics-tames-big-data. Long-term weather forecasters also use big data sets that pose no threats to privacy. Todd Woody, "Meet the Scientists Mining Big Data to Predict the Weather," FORBES, APR. 9, 2012, *available at* http://www forbes.com/sites/toddwoody/2012/03/21/meet-the-scientists-mining-big-data-to-predict-the-weather/.

[10] *See* Big Data Report at 54; RFC, 79 Fed Reg. at 32715.

can adapt the Fair Information Practice Principles (FIPPs) framework to work with big data.[11] These issues need further research and careful consideration by stakeholders. Third, there are concerns over fairness and discrimination in big data. While these are not privacy issues as such, they are important and worth studying carefully. I would like to explore these three points more fully.

Many concerns over big data are not unique to big data. Many big data concerns are also concerns for traditional "small data" and are already familiar to the FTC. For example, without adequate security safeguards, any data, big or small, can fall into the wrong hands. Recent reports of data breaches at retailers and other businesses obviously raise serious concerns. Yet there are real market and reputational incentives for companies to get data security right in the big data context. Furthermore, the FTC has for years been actively enforcing basic data security requirements to address consumer harm and has brought more than fifty-seven data security cases. Recently, in *FTC v. Wyndham*, a federal district court confirmed that the FTC has authority to protect consumers from unfair data security practices by bringing such cases.[12] And while some recent "big data" breaches are very large in scale, this is not a new development. For example, in 2009 the FTC investigated a data breach at Heartland Payment Systems, where

---

[11] *See* DEPT. OF COMMERCE INTERNET POLICY TASK FORCE, *Commercial Data Privacy and Innovation in the Internet Economy: a Dynamic Policy Framework*, at 11 (2010) (describing the 1973 origin of the Fair Information Practice Principles framework at the Department of Health, Education, and Welfare), *available at* http://www ntia.doc.gov/files/ntia/publications/iptf_privacy_greenpaper_12162010.pdf.

[12] *FTC v. Wyndham Worldwide Corp., et al.*, No. 13-1887, 2014 U.S. Dist. LEXIS 47622 (D.N.J. Apr. 7, 2014), *appeal pending*, No. 14-8091 (3d Cir. July 29, 2014).

hackers stole more than 130 million credit card numbers.[13] Big data technology does not raise fundamentally new data security issues.

Similarly, some groups also argue that certain types of particularly sensitive data, such as data about children, health, or finances, deserve heightened protection when stored in big data sets. Of course, the FTC already recognizes the need to more thoroughly protect such types of data, whether the data is in big data or small data environments,[14] and current laws and FTC enforcement reflect that policy.[15]

Tension with Certain Fair Information Practice Principles. Other concerns about big data do appear to raise new issues, however. In particular, maximizing the benefits of big data may create tension with the notice and the purpose limitation principles in the FIPPs. These two related principles say that an information collector should inform consumers about the collection and its purpose and get the consumer's consent for the collection for that purpose. Yet much of the promise of big data is that it can find something new and useful in the data that could not have been anticipated at the time of collection. But companies cannot give notice at the time of collection for unanticipated uses. Furthermore, in many cases, data scientists create one big data set from many other smaller collections that initially served different purposes and may have been collected at different times from a wide range of sources. As such, it is difficult or impossible to notify individuals of the new purpose for which the data is being used.

---

[13] Robert McMillan, "SEC, FTC investigating Heartland after data theft," Feb. 25, 2009, COMPUTERWORLD, *available at* http://www.computerworld.com/s/article/9128658/SEC_FTC_investigating_Heartland_after_data_theft.

[14] FTC REPORT, Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers, FTC Mar. 2012 at 15-16.

[15] *See* Children's Online Privacy Protection Act of 1998, 15 U.S.C. §§ 6501-6506 (COPPA).

The FIPPs principle of data minimization is also in tension with the incentives of big data. Part of the promise of big data is to pull knowledge from data points whose value was previously unknown. Thus, retention of as much data as possible for lengthy amounts of time is a common practice. Strictly limiting the collection of data to the particular task currently at hand and disposing of it afterwards would handicap the data scientist's ability to find new information to address future tasks. Certain de-identification techniques such as anonymization, although not perfect, can help mitigate some of the risks of comprehensive data retention while permitting innovative big data analysis to proceed. I believe FIPPs remains a solid framework and is flexible enough to accommodate a robust big data industry, but we have some work to do to resolve these tensions.

Other, Non-Privacy Concerns. Finally, some advocates worry that companies will use big data techniques to prejudge or discriminate against individuals unfairly or erroneously without recourse. One concern is that a researcher could collect non-sensitive information about a consumer and then use big data analysis to infer certain sensitive characteristics about that consumer.

This is a complicated issue that we need to know more about. Companies have long engaged in this type of consumer targeting with more traditional data tools. It is not clear how to weigh the additional value or risk big data analysis will bring to the long-established practice of targeted advertising.

Second, if companies do engage in this sort of analysis, we need to determine how they might use such information. How data is used matters, as our existing legal framework restricts certain uses of data regardless of how it was collected. Specifically, the Fair Credit Reporting

Act[16] establishes constraints for companies that make certain uses of data provided by "credit reporting agencies": creditworthiness, insurance eligibility, evaluation for employment, and renter background checks. Passed in 1970 in response to the creation of credit reporting bureaus, the FCRA could be considered the first "big data" bill. In fact, the FTC has applied the FCRA in a "big data" context. In 2012, the FTC entered into an $800,000 settlement with Spokeo, a company that assembles personal profiles of individuals from information in public records, white pages, and social networking sites.[17] Spokeo was allegedly marketing personal information to potential employers in violation of the FCRA.

The RFC asks, "Should a responsible use framework, as articulated in Chapter 5 of the Big Data report, be used to address some of the challenges posed by big data?"[18] I believe that it is important to seek an approach that protects consumers from substantial privacy harms while not hampering the economic and societal benefits that data-driven technology may offer. In pursuit of this goal, some have suggested focusing on the use of personal information and the impact on the individual, rather than attempting to safeguard privacy primarily by controlling information collection. Such use-focused approaches emphasize the difficulty of specifying unforeseen but valuable subsequent uses of data. To address these challenges, such approaches offer (in various formulations) a framework that emphasizes preventing harmful uses of personal

---

[16] The Fair Credit Reporting Act, 15 U.S.C. § 1681 et. seq., was enacted in 1970 and significantly amended in 1996, 2003, and 2010. The FCRA does not limit what information may be collected by credit reporting agencies but rather focuses on limiting third party access to credit data for permissible purposes (which do not include marketing), ensuring accuracy of such data, providing consumers notice of adverse actions taken against them based on such data, and ensuring consumer access to and ability to correct data about themselves.

[17] Press Release, FED. TRADE COMM'N, *Spokeo to Pay $800,000 to Settle FTC Charges Company Allegedly Marketed Information to Employers and Recruiters in Violation of FCRA* (June 12, 2012) http://www ftc.gov/news-events/press-releases/2012/06/spokeo-pay-800000-settle-ftc-charges-company-allegedly-marketed.

[18] RFC, 79 Fed Reg. at 32715.

information, accountability for use of personal data however collected, a respect for context, and transparency about the use of such data with a concomitant ability of consumers to know if data has been used to disadvantage them.

I believe it would be useful to explore whether such frameworks, by specifically prohibiting certain clearly impermissible uses of data, could help protect consumers while enabling continued innovation in big data.

None of this is to denigrate the establishment of principles to guide the collection of data. Such principles can and do serve as important best practices or industry standards. Thus, I have supported as best practices many (although not all) of the recommendations of the FTC's 2012 report on "Protecting Consumer Privacy in an Era of Rapid Change."[19] Some of the most relevant recommendations of that report for big data include:

- **Privacy by Design** – Companies should build in consumer privacy protections at every stage in developing their products. These protections include reasonable security for consumer data and reasonable procedures to promote data accuracy. In the big data context, built-in de-identification measures could play an important role in protecting consumer privacy.
- **Simplified Choice for Businesses and Consumers** – Recognizing that there is no single best way to offer notice and choice in all circumstances, companies should adopt notice and choice options that appropriately reflect the context of the transaction or the relationship the company has with the consumer. In the big data context, this may be challenging, but I believe it is a principle worth continuing to pursue.
- **Greater Transparency** – Companies should disclose details about their collection and use of consumers' information and provide consumers access to the data collected about them.

I believe these best practices are flexible enough to remain useful in many, if not all, situations. Companies that embrace these principles would benefit their customers. Of course,

---

[19] FED. TRADE COMM'N, *Protecting Consumer Privacy in an Era of Rapid Change: Recommendations For Businesses and Policymakers* (Mar. 2012) http://www ftc.gov/reports/protecting-consumer-privacy-era-rapid-change-recommendations-businesses-policymakers.

best practices necessarily change with the environment. Policy makers and stakeholders must work together to determine what changes in best practices may be necessary to protect and advance consumer welfare.

III. FTC's Role in Big Data

The FTC is working to ensure that the promise of big data is realized by using our unique set of enforcement and policy tools. First, the FTC is an enforcement agency and it can and should use its traditional deception and unfairness authority to stop consumer harms that may arise from the misuse of big data. The FTC has a strong track record of protecting consumers from harms, whether those harms are inflicted by bad actors using traditional or cutting-edge technological tools. The FTC's action in *Spokeo*, mentioned above, is an example of the agency's policing of harmful big data uses. Strong enforcement in the big data area will help not only consumers but also other companies using big data analysis by policing actors that may tarnish the technology itself.

Second, we are using our convening power and our policy and R&D functions to better understand big data technology; the new business models it may enable; the applicability of existing regulatory structures, including self-regulation; market dynamics; and the nature and extent of likely consumer and competitive benefits and risks. The FTC is currently working to understand the promise and risks of big data and related technologies. Last year the FTC held a workshop on the Internet of Things in which we explored both the potential benefits and risks to consumers of this new environment of constant data flow.[20] More recently, the FTC hosted a

[20] Press Release, FED. TRADE COMM'N, *FTC Announces Agenda, Panelists for Upcoming Internet of Things Workshop* (Nov. 8, 2013) http://www.ftc.gov/news-events/press-releases/2013/11/ftc-announces-agenda-panelists-upcoming-internet-things-workshop.

workshop on alternative scoring mechanisms to evaluate the potential implications of new types of scoring that rely on big data predictive analytics to provide identity verification, fraud prevention, and marketing and other services.[21] In May, the Commission hosted an event on consumer generated and controlled health data – one of the newest and most interesting sources of big data.[22] And this September the FTC will host a big data workshop to explore some of the benefits and potential risks of various big data techniques.[23]

As the FTC and other policy makers engage with big data issues, I believe two principles should guide our work. First, as with all dynamic markets, policy makers must approach big data technologies with what I call regulatory humility. Our most successful technological advances, such as the Internet itself, have generated massive amounts of consumer welfare and have thrived largely because market participants have enjoyed wide latitude to experiment with new technology-driven business models, allowing the market to determine which of those models succeeds or fails. Particularly as we consider whether big data could be used for discrimination, we should look at the lessons of recent technological history: in many ways, information technology has been a great equalizing force, giving voice and the power of information to the individual.

---

[21] Press Release, FED. TRADE COMM'N, *FTC Announces Agenda, Panelists for Alternative Scoring Seminar* (Mar. 14, 2014) http://www ftc.gov/news-events/press-releases/2014/03/ftc-announces-agenda-panelists-alternative-scoring-seminar.

[22] Press Release, FED. TRADE COMM'N, F*TC to Host Spring Seminars on Emerging Consumer Privacy Issues* (Dec. 2, 2013) http://www ftc.gov/news-events/press-releases/2013/12/ftc-host-spring-seminars-emerging-consumer-privacy-issues.

[23] Press Release, FED. TRADE COMM'N, *FTC to Examine Effects of Big Data on Low Income and Underserved Consumers at September Workshop* (Apr. 11, 2014) http://www.ftc.gov/news-events/press-releases/2014/04/ftc-examine-effects-big-data-low-income-underserved-consumers.

Second, policy makers must identify substantial consumer harm before taking action. The FTC in particular should remain vigilant for deceptive and unfair uses of big data, but should avoid preemptive action that could preclude entire future industries. Ultimately, our work as an agency should help strengthen competition and the market to better provide beneficial outcomes in response to consumer demand, rather than to try to dictate desired outcomes to the market.

I do believe some specific changes are needed to better inform consumers about big data. The FTC recently recommended in our Data Broker Report that Congress consider requiring data brokers to give consumers more transparency and control over the information such brokers have about individual consumers.[24] I believe a proper implementation of this recommendation could help consumers embrace the big data movement by revealing the value of such information and reducing the fear of the unknown.

IV. Conclusion

To conclude, big data is a powerful tool with great promise and some risks. Many of the concerns raised by big data are suitably handled by current law and policy. But where there are new issues, regulators need to work with innovators to understand the issues deeply and focus our enforcement actions on situations where improper use of consumer information causes substantial harm. This approach will free entrepreneurs to innovate with big data tools while simultaneously helping to ensure that consumer privacy remains protected. I applaud NTIA's

[24] FED. TRADE COMM'N, *Data Brokers: A Call for Transparency and Accountability*, 49 (May 2014), http://www ftc.gov/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014.

efforts to explore these issues and look forward to future opportunities to work together to ensure consumers benefit from big data technology while mitigating any risks.